Sleep Health & Lifestyle Prediction

Python ML Project

Project Roadmap

Objective

Predict whether a person has a sleep disorder (insomnia, apnea, or none) using **health and lifestyle data** such as sleep duration, stress level, work hours, BMI, and daily habits. Goal: provide actionable insights into lifestyle factors affecting sleep health.

Tools & Technologies

- **Python** – Core programming
- **Pandas** – Data collection, cleaning, and preprocessing
- **Matplotlib & Seaborn** – Data visualization
- **Scikit-learn** – Machine learning model building

Process Workflow

1. **Data Collection & Exploration**
   - Collected dataset containing sleep duration, stress levels, work hours, BMI, and activity.
   - Explored data distribution and relationships using plots and charts.
2. **Data Preprocessing**
   - Handled missing values.
   - Encoded categorical features using Label Encoding.
   - Scaled numerical features using StandardScaler for uniformity.
3. **Train-Test Split**
   - Divided dataset into **training (80%)** and **testing (20%)** sets using `train_test_split`.
4. **Model Training**
   - Trained a **Random Forest Classifier** for prediction.
   - Tuned parameters for improved accuracy.
5. **Model Evaluation**
   - Measured performance using **Accuracy, Classification Report, and Confusion Matrix**.
   - Visualized confusion matrix heatmap for insights.
6. **Feature Importance Analysis**
   - Identified key factors affecting sleep disorders: **stress level, work hours, sleep duration**.

Outcome

- Developed a **reliable prediction model** for sleep disorders.
- Discovered that lifestyle factors like stress, long work hours, and insufficient sleep are the most influential.
- Demonstrates how **ML with health data** can guide lifestyle improvements.
- Can be extended for larger datasets or real-time monitoring.